

IN THE UNITED STATES  
PATENT AND TRADEMARK OFFICE

**PATENT APPLICATION**

Applicant: **BEERENDS, John Gerard**

Case: **PTT-133(402582US)**

International Application No.: **PCT/EP00/08884**

International Filing Date: **08 September 2000**

Priority Date Claimed: **15 September 1999**

Title: **MEASURING THE PERCEPTUAL QUALITY OF SPEECH SIGNALS  
INCLUDING ECHO DISTURBANCES**

COMMISSIONER FOR PATENTS  
**BOX PCT**  
Washington, D. C. 20231

S I R:

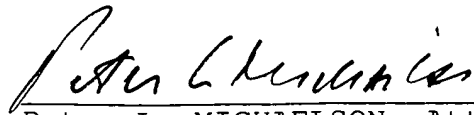
**SUBMISSION OF PRIORITY DOCUMENT**

In connection with the above-captioned application, applicant encloses the following certified priority document (with English translation) to support the claim to priority:

Netherlands - Serial No. 1013044, filed  
September 15, 1999.

Respectfully submitted,

25 February 2002

  
\_\_\_\_\_  
Peter L. MICHAELSON, Attorney  
Reg. No. 30,090  
Customer No. 007265  
(732) 530-6671

**THIS PAGE BLANK (USPTO)**

10/069704

Rec'd PCT/PTO 26 FEB 2002


MICHAELSON & WALLACE  
Counselors at Law  
Parkway 109 Office Center  
328 Newman Springs Road  
P.O. Box 8489  
Red Bank, New Jersey 07701

**\*\*\*EXPRESS MAIL CERTIFICATION\*\*\***

"Express Mail" mailing label number: **EL632364666US**

Date of deposit: **26 February 2002**

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner for Patents, **Box PCT**, Washington, D.C. 20231.

  
\_\_\_\_\_  
Signature of person making certification

Peter L. MICHAELSON  
\_\_\_\_\_  
Name of person making certification

**THIS PAGE BLANK (USPTO)**

10/069704

REC'D 03 NOV 2000

WIPO

PCT

KONINKRIJK DER



NEDERLANDEN

EP00/8884

Bureau voor de Industriële Eigendom



BEST AVAILABLE COPY

4

Hierbij wordt verklaard, dat in Nederland op 15 september 1999 onder nummer 1013044,  
ten name van:

**KONINKLIJKE KPN N.V.**

te Groningen

een aanvraag om octrooi werd ingediend voor:

"Kwaliteitsmeting van spraaksignalen met echo-verstoring",

en dat de hieraan gehechte stukken overeenstemmen met de oorspronkelijk ingediende stukken.

**PRIORITY DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(A) OR (B)

Rijswijk, 26 juni 2000.

De Directeur van het Bureau voor de Industriële Eigendom,  
voor deze,

A.W. van der Kruk

Uittreksel

Voor het meten van de invloed van echo op de perceptiekwaliteit aan de sprekerszijde van een telefoonverbinding in een telecommunicatienetwerk wordt het spraaksignaal ( $s(t)$ ) van de spreker en een gecombineerd signaal ( $s'(t)$ ) toegevoerd aan een objectieve meetinrichting (32), zoals een PSQM-systeem, voor het verkrijgen van een uitgangssignaal ( $q(t)$ ) dat een schattingswaarde representeert van de perceptiekwaliteit. Het gecombineerde signaal wordt verkregen in een signaalcombinator (34) door signaalcombinatie van een met het spraaksignaal corresponderend, uit het netwerk afkomstig echo-signaal ( $e(t)$ ) en het spraaksignaal ( $s(t)$ ). De meetwijze is van toepassing op in een database vastgelegde spraaksignalen, maar ook op rechtstreeks uit een telefoonverbindingscircuit verkregen signalen, waarbij het uitgangssignaal wordt toegevoerd aan een kwaliteitsbewakingssysteem en/of aan een echo-minimaliserende inrichting.

(FIG. 3)

Titel: Kwaliteitsmeting van spraaksignalen met echo-verstoring

A. Achtergrond van de uitvinding

1. Gebied van de uitvinding

De uitvinding ligt op het gebied van het meten van de conversationele kwaliteit van telecommunicatiesystemen. Meer in het bijzonder betreft zij het meten van de invloed van echo-verstoringen op de subjectief waargenomen conversationele kwaliteit van een telefoonverbinding in een telecommunicatiesysteem.

2. Stand van de techniek

Echo is een significante factor in de perceptie van de gesprekskwaliteit in een "end-to-end" telefoonverbinding. De invloed van deze factor op de perceptie kan worden gekwantificeerd door meting van de combinatie van een tweetal parameters die bekend zijn onder de aanduidingen "talker echo loudness rating" en "round-trip delay", zoals bijvoorbeeld gedefinieerd in de Recommendation G.107 van ITU-T (Referentie [1]; voor meer bibliografische details met betrekking tot de referenties, zie hierna onder D.). Voor een niet-verstorende meting ("non-intrusive measurement") voor het bepalen van dergelijke parameters geeft ITU-T Recommendation P.561 (zie Referentie [2]) aanbevelingen. Deze bekende techniek heeft echter de volgende nadelen. Om de invloed van de echo-verstoring te meten is allereerst een correlatie-techniek nodig om de vertraging ("delay") van de echo te bepalen. Voorts is een specifiek algoritme nodig voor het schatten van de "talker echo loudness rating". Verder is het bepalen van de echo-vertraging moeilijk wanneer de echo-signalen een laag niveau bezitten, i.e. wanneer zij zich nabij de waarneembaarheidsgrens bevinden, zoals veelal het geval is op spraakverbindingen van hoge-kwaliteit. Zelfs al wordt van de vertraging en de luidheid een correcte schatting gemaakt, dan nog kunnen de verschillen in perceptie van de echo-verstoring tussen verschillende sprekers niet worden gemeten, aangezien bij de meting geen details van het spraaksignaal in rekening kunnen worden gebracht. Derhalve vertoont de bekende techniek slechts een matige correlatie tussen de objectieve meetresultaten en de meer subjectieve bevindingen van de sprekers, vooral bij geringe echo-verstoringen.

In ditte moet bij echo op een telefoonverbinding tussen een A-abonnee en een B-abonnee onderscheid worden gemaakt tussen zogeheten spreker-echo ("talker echo") (van A naar A) die de spreker aan de mondside (A) kan ervaren tijdens het spreken, en luisteraar-echo ("listener echo") die de luisteraar aan de ontvangerside (B) tijdens het luisteren kan ervaren. De perceptie van de invloed van de luisteraar-echo op de spraakverbinding kan worden verbeterd door de spreker te informeren over de aanwezigheid van de echo.

spraaksignalen, kan in principe worden gekwantificeerd met behulp van een zogeheten objectieve meettechniek, zoals de Perceptual Speech Quality Measure (PSQM). Met deze bekende techniek (zie bijvoorbeeld referentie [3] en [4]), die de perceptie-eigenschappen van het menselijke gehoor modelleert, kan de kwaliteit van een spraakverbinding in één richting ("one-way") worden voorspeld. Het kwantificeren van de invloed van echo op de perceptie van de spreker met deze techniek is echter niet bekend en niet zonder meer mogelijk, maar wel gewenst.

B. Samenvatting van de uitvinding

De uitvinding beoogt te voorzien in een werkwijze en een inrichting voor het meten van de invloed van echo op de perceptiekwaliteit aan de sprekerszijde van een telefoonverbinding in een telecommunicatienetwerk, die en genoemde nadelen van de bekende techniek niet bezit, en tegemoet komt aan genoemde wens.

Voorts beoogt zij te voorzien in een telefoonverbindingscircuit waarin de werkwijze en de inrichting wordt toegepast.

Een werkwijze en een inrichting voor het meten van de conversationele kwaliteit van een telefoonverbinding in een telecommunicatienetwerk volgens de aanhef van conclusie 1, en conclusie 9, respectievelijk, voor de definitie waarvan referentie [4] is aangewend, hebben volgens de uitvinding respectievelijk het kenmerk van conclusie 1 en van conclusie 9.

In een voorkeursuitvoering van de uitvinding heeft de werkwijze het kenmerk van conclusie 1, en de inrichting het kenmerk van conclusie 10.

De uitvinding is gebaseerd op het inzicht dat een sprekende telefoongebruiker tegelijkertijd luistert en derhalve zijn eigen spraaksignaal tegelijkertijd hoort met een eventueel uit de hoorn van het telefoontoestel terugkomende echo van zijn spreken. Derhalve zal bij toepassing van een oorspronkelijk spraaksignaal en een gecombineerd signaal samengesteld uit het oorspronkelijke spraaksignaal en een corresponderend echo-signaal als ingangssignalen voor een objectieve meettechniek zoals PSQM wel tot een bruikbare schatting van de perceptiekwaliteit kunnen leiden, terwijl dit niet het geval is als alleen het oorspronkelijke spraaksignaal en het corresponderende echo-signaal worden gebruikt.

Een telefoonverbinding is te voorzien voor een telefoonverbinding in een telecommunicatienetwerk, bestaande uit een heen-kanal en een retour-kanal, en een tussen het heen-kanal en het retour-kanal opgenomen echo-minimaliserende inrichting, die het oorspronkelijke spraaksignaal en het corresponderende echo-signaal als ingangssignalen voor een objectieve meettechniek zoals PSQM kan gebruiken.



Verdere voorkeursuitvoeringen van de uitvinding zijn samengevat in verdere onderconclusies.

C. Referenties

- [1] ITU-T Recommendation G.107: The E-model, a computational model for use in transmission planning, December 1998;
- [2] ITU-T Recommendation P.561: In-service, non-intrusive measurement device-voice service measurements, February 1996;
- [3] J.B. Beerends and J. A. Stemerdink, A perceptual speech quality measure based on a psychoacoustic sound representation. J. Audio Eng. Soc. 42:115-123, March 1994;
- [4] ITU-T Recommendation P.861: Objective quality measurement of telephoneband (300-3400 Hz) speech codecs, August 1996.

Alle referenties worden beschouwd als geïncorporeerd in de onderhavige aanvraag.

D. Korte beschrijving van de tekening

De uitvinding zal nader worden toegelicht middels een beschrijving van uitvoeringsvoorbeelden, waarbij wordt verwezen naar een tekening die de volgende figuren omvat:

- FIG. 1 toont schematisch een bekende werkwijze voor het meten van de perceptiekwiteit van een spraaksignaal;
- FIG. 2 toont schematisch een telefoonverbinding in een telecommunicatienetwerk;
- FIG. 3 toont schematisch een werkwijze volgens de uitvinding;
- FIG. 4 toont een deel van een telefoonverbindingscircuit waarin de uitvinding wordt toegepast.

E. Beschrijving van uitvoeringsvoorbeelden

FIG. 1 toont een bekende schematische opzet van een toepassing van een objectieve meettechniek zoals bijvoorbeeld die gebaseerd op een model van het menselijk gehoor en die wel wordt aangeduid met PQM, voor het meten van de perceptiekwiteit van spreekverbindingen. Zij omvat een spreker 1, telecommunicatienetwerk onder test 10, hierna te beschrijven aangeduid door netwerk 10, en een systeem 12 voor de perceptuele analyse van aangeboden spraaksignalen, hierna slechts schematische aangeduid door PQM-systeem 12. Een spraaksignaal dat wordt gebruikt enerzijds als inversignaal van het netwerk 10, en anderzijds als een eerste inversignaal van het PQM-systeem 12. Een tweede inversignaal wordt van het netwerk 10, in de tekening ook het tweede inversignaal, afgeleid van het eerste inversignaal, wat kan worden afgeleid van het eerste inversignaal.

invoersignaal van het PSQM-systeem 12. Een uitvoersignaal  $p(t)$  van het PSQM-systeem 12 representeert een schatting van de perceptiekwaliteit van de spreekverbinding door het netwerk 10. Omdat de invoerszijde en de uitvoerszijde van een spreekverbinding, vooral in geval deze door een telecommunicatienetwerk voert, zich op afstand bevinden, wordt voor de invoersignalen van het PSQM-systeem meestal gebruik gemaakt van op databases opgeslagen spraaksignalen. Hierbij wordt zoals gebruikelijk onder spraaksignaal verstaan elk in principe voor het menselijke gehoor waarneembaar geluid zoals spraak en tonen. Het systeem of netwerk onder test kan vanzelfsprekend ook een simulatiesysteem zijn, dat een telecommunicatienetwerk simuleert. Met deze bekende techniek zijn betrouwbare schattingen van de perceptiekwaliteit mogelijk.

In FIG. 2 is schematisch een staande telefoonverbinding tussen een A-abonnee en een B-abonnee van een telecommunicatienetwerk 20 weergegeven. Telefoon toestellen 21 en 22, respectievelijk van de A-abonnee en de B-abonnee, zijn via tweedraadsverbindingen 23 en 24 en vierdraads-interfaces, t.w. vorkschakelingen 25 en 26 met het netwerk 20 verbonden. Door het netwerk heeft de staande telefoonverbinding een heen-kanaal 27 waarover spraaksignalen van de A-abonnee worden geleid, en een retour-kanaal 28 waarover spraaksignalen van de B-abonnee worden geleid. Een spraaksignaal  $s(t)$  dat de microfoon van het telefoontoestel 21 van de A-abonnee treft, wordt via het heen-kanaal 27 van de telefoonverbinding doorgegeven naar de bortelefoon van het telefoontoestel 22, en wordt daar hoorbaar voor de B-abonnee als een door het netwerk beïnvloed spraaksignaal  $s'(t)$ . Elk spraaksignaal op het heen-kanaal 27 veroorzaakt in het algemeen, vooral door de aanwezigheid van genoemde vorkschakelingen, een echo-signaal  $e(t)$  op het retour-kanaal 28 van de telefoonverbinding, en dat wordt doorgegeven aan de bortelefoon van het telefoontoestel 21, en derhalve daar de A-abonnee tijdens het spreken kan hinderen. Om dergelijke echo-effecten tot een minimum te beperken is het algemeen gebruikelijk om in de telefoonverbindingscircuits echo-onderdrukkers ("echo-suppressors") of echo-opheffers "echo-cancellors" opgenomen, één aan elk uiteinde van een telefoonverbinding. Een echo-onderdrukker of opheffer maakt continu vergelijking van de signalen aanwezig in het heen- en het retour-kanaal en schakelt het echo-signaal en brengt dit in mindering op het signaal in het heen-kanaal. Een dergelijke schakeling is echter niet noodzakelijk om te voeren, hetgeen vooral het geval is op hoge kwaliteitsniveau's of lijnen, waar de echo-signalen zich op de grens van het hoorbaar gebied bevinden. De schakeling moet hier de binnengekomene echo's niet terugkoppelen naar de afzender, maar moet de afzender van de echo's in kennis stellen. Het is mogelijk dat de afzender de echo's niet

dergelijke optimalisatie-routines althans mede gebaseerd op de hierna onder verwijzing naar FIG. 3 beschreven werkwijze, dan is een optimaal resultaat met een minimaal waarneembare echo mogelijk.

FIG. 3 toont een schematische opzet volgens de uitvinding voor het verkrijgen van een schatting van de perceptiekwaliteit van een telefoonverbinding voor een telefoongebruiker tijdens zijn spreken op zijn eigen telefoontoestel. Op soortgelijke wijze als de opzet van FIG. 1 omvat de opzet van FIG. 3 een systeem of telecommunicatienetwerk onder test 30, hierna kortweg netwerk 30 genoemd, en een systeem 32 voor de perceptuele analyse van aangeboden spraaksignalen, hierna slechts kortheidshalve aangeduid door PSQM-systeem 32. Een willekeurig spraaksignaal  $s(t)$  wordt enerzijds gebruikt als invoersignaal van het netwerk 30, en anderzijds als eerste invoersignaal van het PSQM-systeem 32. Een uit het netwerk 30 verkregen echosignaal  $e(t)$ , dat correspondeert met het ingevoerde spraaksignaal  $s(t)$ , wordt in een combinatiecircuit 34 gecombineerd met het spraaksignaal  $s(t)$  tot een gecombineerd spraaksignaal  $s'(t)$ , dat vervolgens als tweede invoersignaal van het PSQM-systeem wordt gebruikt. Een uitvoersignaal  $q(t)$  van het PSQM-systeem 32 representeert een schatting van de perceptiekwaliteit van de spreekverbinding door het netwerk 30 zoals ervaren door de telefoongebruiker tijdens het spreken op zijn eigen toestel. Ook hier kan weer gebruik gemaakt worden van op databases opgeslagen signalen. Deze kunnen bijvoorbeeld zijn verkregen uit het telefoontoestel (respectievelijk door elektrische uitkoppeling uit het microfoon-circuit en het oortelefooncircuit van de A-abonnee bij een staande verbinding tijdens gespreksstilte van de B-abonnee. Aangezien echter de tweedraadsverbinding tussen telefoonaansluiting bij de abonnee en de vierdraads-interface met het netwerk niet of nauwelijks bijdraagt aan het echo-signaal  $e(t)$  (vanzelfsprekend wel aan het echo-signaal dat optreedt bij de B-abonnee), kunnen de signalen  $s(t)$  en  $e(t)$  ook worden afgekoppeld uit het heenkanaal en het retourkanaal nabij de vierdraads-interface. Dit biedt de mogelijkheid van een permanente meting van de perceptiekwaliteit bij staande telefoonverbindingen. FIG. 4 toont hieraan schematisch een uitvoeringsvorm.

FIG. 4 toont op soortgelijke wijze als een deel van FIG. 2 een tweedraadslijn 41 die via een vierdraadsinterface, i.e. afschakeling 42 met een telecommunicatienetwerk 40 is verbonden en via welke een abonnee te spreken verbonden met een heenkanaal 43 en een retourkanaal 44 door het netwerk kan worden opgehoord. In het bij de tweedraadslijn 41 behorende lijncircuit is een echo-punt 46 aangebracht. Dit kan bijvoorbeeld een PSQM-systeem 32, waarvan een uitvoersignaal  $q(t)$  de perceptiekwaliteit van de verbinding tussen de abonnee en het netwerk 40, wordt gemeten.

ingangspoort 48 is gekoppeld met een uitgangspoort van een signaalcombinator 49 met twee ingangspoorten die respectievelijk zijn gekoppeld met het heen-kanaal 44 en het retour-kanaal 45. Een uitgangspoort 50 van het PSQM-systeem 42 kan voor

5 kwaliteitsbewakingsdoeleinden direct of via een schakelaar 51 gekoppeld zijn (pijlen F en G) met een bewakingssysteem (niet getekend). Ook kan de uitgangspoort 50, zoals getekend, via de schakelaar 51 zijn gekoppeld met een sturingang 52 van de echo-opheffer 46. De schakelaar 51 wordt bij voorkeur bestuurd door een stuursignaal afgegeven door een  
10 detectiecircuit 53 (bijvoorbeeld uitgevoerd als een op zich bekend "double-talk" detectiecircuit), dat is gekoppeld met het retourkanaal, voor het detecteren van de gesprekstoestand op het retourkanaal 45, zoals bijvoorbeeld gespreksstilte aan de zijde van een B-abonnee. Het via de uitgangspoort 50 van het PSQM-systeem beschikbaar komende  
15 schattingssignaal kan aldus enerzijds worden gebruikt voor allerlei kwaliteitsbewakingsdoeleinden, en anderzijds direct worden gebruikt in echo-minimaliserende apparatuur.

In de meest eenvoudige uitvoeringen zijn het combinatiecircuit 34 en de signaalcombinator 49 signaaloptellers. Bij het in de praktijk toepassen van de methode en de inrichting wordt in de signaalcombinator het  
20 uitvoeren van de optelfunctie (additie) bij voorkeur voorafgegaan door het zogencemd "invers-filteren" van een van de signaalcomponenten. Het daarbij toegepaste inverse-filter genereert een lineaire schatting van het echo-pad en draagt in belangrijke mate bij aan het bereiken van een  
25 hoge correlatie tussen een objectieve meting en een subjectieve waarneming.

F. Conclusies

1. Een werkwijze voor het meten van de conversationele kwaliteit van een telefoonverbinding in een telecommunicatienetwerk, omvattende de volgende stappen:

- het ontvangen van een aan een zendszijde van de telefoonverbinding te verzenden of verzonden eerste spraaksignaal,
- het ontvangen van een tweede spraaksignaal dat een functie is van het in het telecommunicatienetwerk beïnvloed eerste spraaksignaal,
- het met behulp van een objectieve meettechniek verwerken van het eerste en het tweede spraaksignaal, waarbij een uitgangssignaal wordt geproduceerd dat een schattingswaarde met betrekking tot de conversationele kwaliteit representeert, met het kenmerk, dat het tweede spraaksignaal wordt gegenereerd door signaalcombinatie van een aan de zendszijde van de telefoonverbinding tijdens het verzenden van het eerste spraaksignaal afgenomen echosignaal en het eerste spraaksignaal.

2. Werkwijze volgens conclusie 1 met het kenmerk, dat de signaalcombinatie een signaaladditie omvat van het echosignaal en het eerste spraaksignaal.

3. Werkwijze volgens conclusie 1 met het kenmerk, dat de signaalcombinatie een signaaladditie omvat voorafgegaan door een inverse filtering van hetzij het echosignaal, hetzij het eerste spraaksignaal.

4. Werkwijze volgens conclusie 1, 2 of 3 met het kenmerk, dat het echosignaal afkomstig is van een retourkanaal van de telefoonverbinding.

5. Werkwijze volgens conclusie 1, 2, 3 of 4 met het kenmerk, dat het eerste spraaksignaal, en hetzij het tweede spraaksignaal, hetzij het echosignaal in een database vastgelegde signalen zijn.

6. Werkwijze volgens conclusie 1, 2, 3 of 4 met het kenmerk, dat het eerste spraaksignaal en het echosignaal in een bundel te worden afgevoerd respectievelijk van een bronkanaal en een retourkanaal van een telefoonverbinding.

7. Werkwijze volgens conclusie 1 met het kenmerk, dat het eerste spraaksignaal en het echosignaal in een bundel te worden afgevoerd van een bronkanaal en een retourkanaal van een telefoonverbinding.

sturingang van een in de staande telefoonverbinding opgenomen echo-minimaliserende inrichting.

8. Werkwijze volgens conclusie 6 of 7 met het kenmerk, dat het  
5 uitgangssignaal van de objectieve meting wordt toegevoerd aan een  
bewakingssysteem.

9. Een inrichting voor het meten van de conversationele kwaliteit  
van een telefoonverbinding in een telecommunicatienetwerk, omvattende:

10 - een eerste ingangspoort voor het ontvangen van een aan een  
zendzijde van de telefoonverbinding te verzenden of verzonden eerste  
spraaksignaal,  
- een tweede ingangspoort voor het ontvangen van een tweede  
15 spraaksignaal dat een functie is van het in het telecommunicatienetwerk  
beïnvloed eerste spraaksignaal,  
- een uitgangspoort voor een uitgangssignaal, en  
- verwerkingsmiddelen voor het met behulp van een objectieve  
meettechniek verwerken van de via de ingangspoorten ontvangen eerste en  
20 tweede spraaksignalen, waarbij het uitgangssignaal wordt geproduceerd,  
dat een schattingswaarde met betrekking tot de conversationele kwaliteit  
representeert,  
met het kenmerk, dat  
de inrichting voorts omvat signaalcombinatiemiddelen voor het combineren  
van een aan de zendzijde van de telefoonverbinding tijdens het verzenden  
25 van het eerste spraaksignaal afgenomen echosignaal en het eerste  
spraaksignaal, en het genereren van het tweede spraaksignaal.

10. Inrichting volgens conclusie 9 met het kenmerk, dat de  
signaalcombinatie-middelen een signaalopteller omvatten.

11. Inrichting volgens conclusie 9 of 10 met het kenmerk, dat de  
eerste en de tweede ingangspoorten zijn gekoppeld met een database van  
spraaksignalen, waarop het eerste spraaksignaal, en hetzij het tweede  
30 spraaksignaal, hetzij het echosignaal zijn vastgelegd.

12. Inrichting volgens conclusie 9, 10 of 11 met het kenmerk, dat de  
signaalcombinatiemiddelen zijn voorzien van een eerste en een tweede  
signaalinput die respectievelijk zijn gekoppeld met het hoorkanaal en  
het zendkanaal van een staande telefoonverbinding, en dat de eerste  
ingangspoort is gekoppeld met het hoorkanaal en de tweede ingangspoort  
35 is gekoppeld met het zendkanaal van de staande telefoonverbinding.

13. Inrichting volgens conclusie 12 met het kenmerk, dat de uitgangspoort is gekoppeld met een sturingang van een in de staande telefoonverbinding opgenomen echo-minimaliserende inrichting.

5 14. Telefoonverbindingscircuit voor een telefoonverbinding in een telecommunicatienetwerk, omvattende een heen-kanaal en een retour-kanaal, en een tussen het heen-kanaal en het retour-kanaal opgenomen echo-minimaliserende inrichting,

10 met het kenmerk, dat het telefoonverbindingscircuit voorts omvat  
 - een objectieve meetinrichting voorzien van een eerste en een tweede ingangspoort, en een uitgangspoort, voor het verwerken van een op de eerste ingangspoort ontvangen eerste spraaksignaal en van een op de tweede ingangspoort ontvangen tweede spraaksignaal, en voor het produceren van een uitgangssignaal op de uitgangspoort, dat een schattingswaarde representeert met betrekking tot de conversationele  
 15 kwaliteit van de telefoonverbinding, waarbij de eerste ingangspoort van de meetinrichting is gekoppeld met het heen-kanaal van de telefoonverbinding, en

20 - een signaalcombinator voorzien van een eerste en een tweede signaalingang welke respectievelijk zijn gekoppeld met het heen-kanaal en het retour-kanaal van de telefoonverbinding, en van een signaaluitgang welke is gekoppeld met de tweede ingangspoort.

15. Telefoonverbindingscircuit volgens conclusie 14 met het kenmerk, dat de uitgangspoort een signaalkoppeling heeft met een bewakingssysteem.

16. Telefoonverbindingscircuit volgens conclusie 14 of 15 met het kenmerk, dat de uitgangspoort van de meetinrichting een signaalkoppeling heeft met een sturingang van de echo-minimaliserende inrichting.

17. Telefoonverbindingscircuit volgens conclusie 14 of 15 met het kenmerk, dat voorts is voorzien in een detectieinrichting voor het detecteren van de besprekingsstand over de staande telefoonverbinding, en in een in de signaalkoppeling opgenomen schakelaar welke wordt aangestuurd door de detectieinrichting.

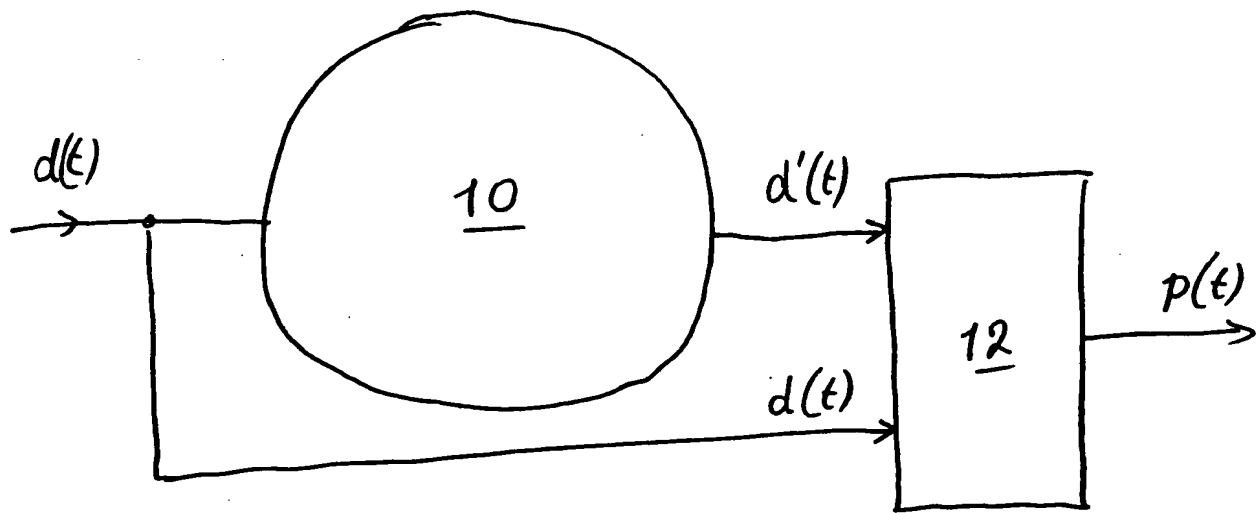


FIG. 1

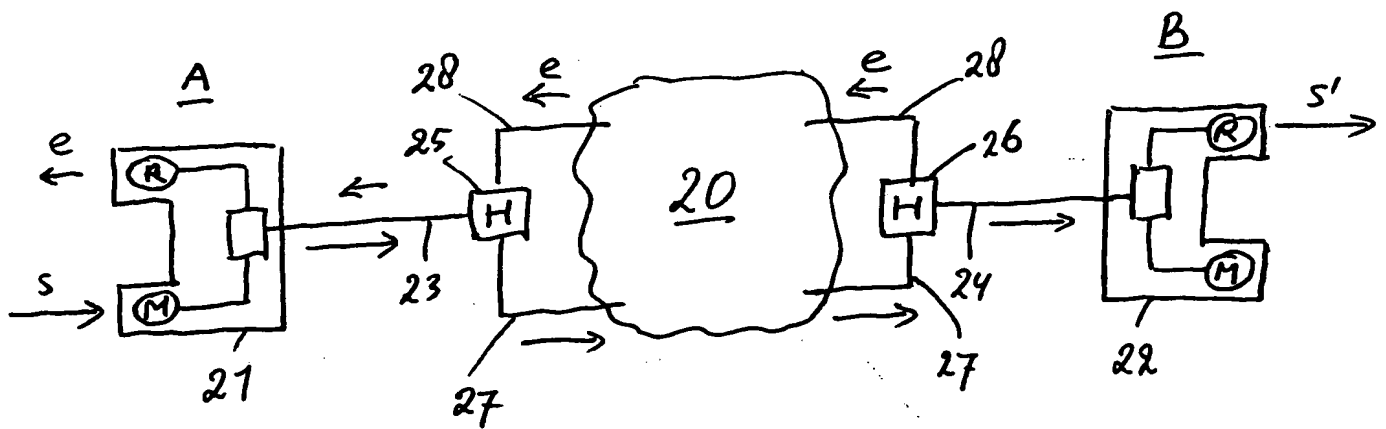


FIG. 2



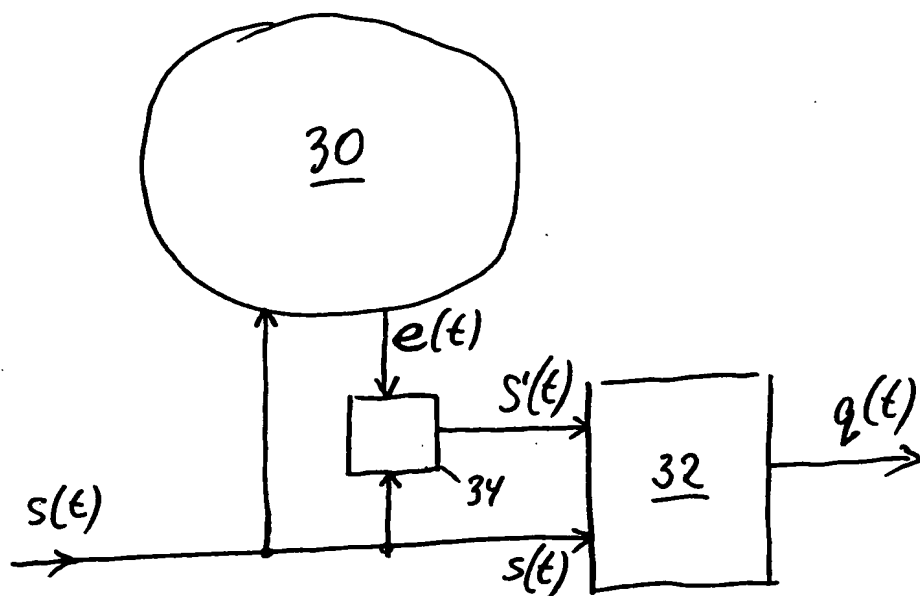


FIG. 3

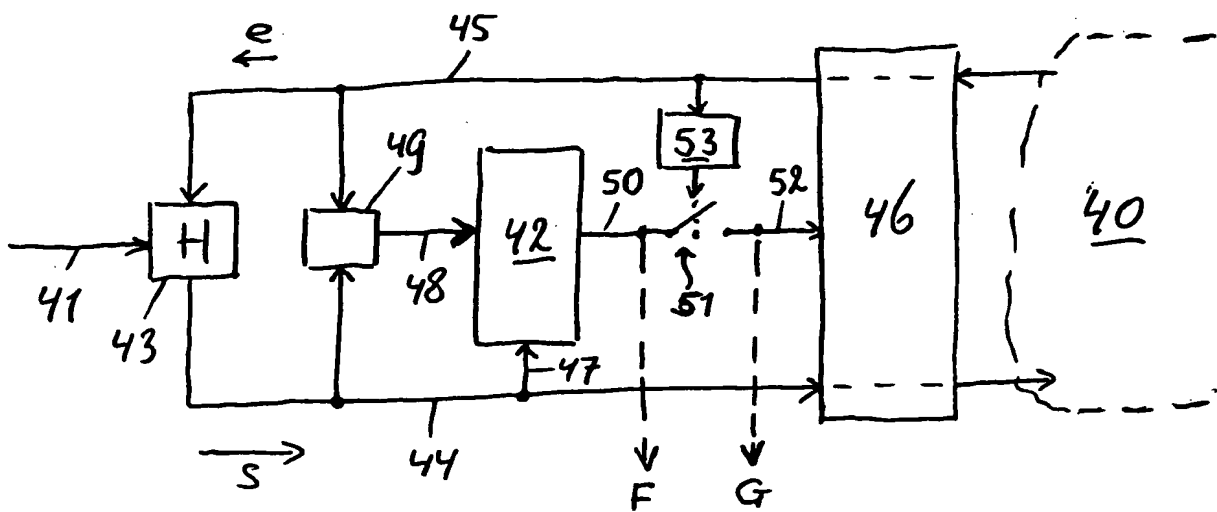


FIG. 4

**This Page Blank (uspto)**

KINGDOM OF THE (crest) NETHERLANDS

PATENT OFFICE

This certifies that in the Netherlands on 15 September 1999 a patent application was filed under number 1013044, in the name of:

**Koninklijke KPN N.V.**

of Groningen

for: "Measuring the quality of speech signals having an echo disturbance."

and that the documents attached hereto are in accordance with the documents originally submitted.

Rijswijk, 26 June 2000

On behalf of the Chairman of the Patent Office,

---

(signature)

(A.W. van der Kruk)

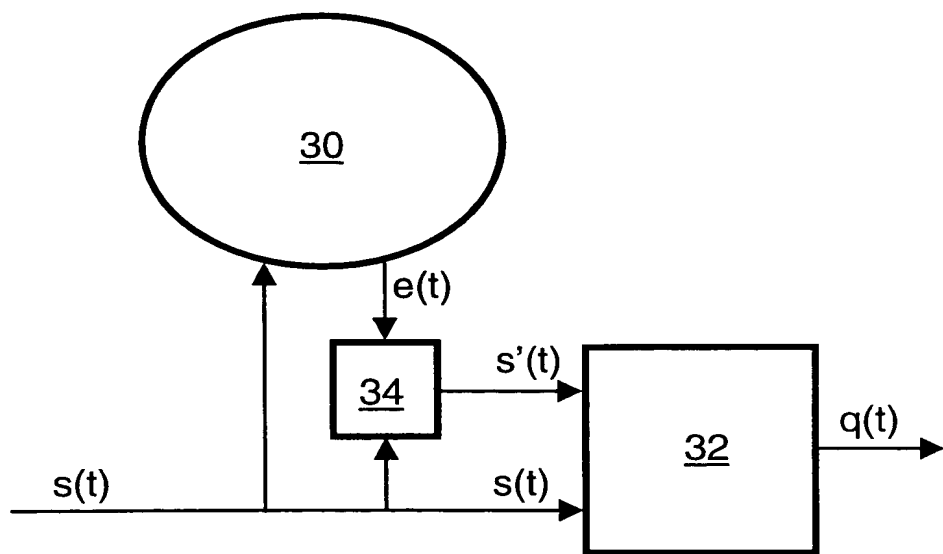
**THIS PAGE BLANK (USPTO)**

# ABSTRACT

For measuring the influence of echo on the perceptual quality on the speaker's side of a telephone link in a telecommunications network, the speech signal ( $s(t)$ ) of the speaker and a combined signal ( $s'(t)$ ) are fed to an objective measurement device (32), such as a PSQM system, for obtaining an output signal ( $q(t)$ ) representing an estimated value of the perceptual quality. The combined signal is obtained in a signal combiner (34) by signal combination of an echo signal ( $e(t)$ ) originating from the network and corresponding to the speech signal, and the speech signal ( $s(t)$ ). The measuring method is applicable to speech signals laid down in a data base, but also to signals directly obtained from a telephone-link circuit, the output signal being fed to a quality-control system and/or to an echo-minimising device.

(FIG. 3)

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**



## Measuring the quality of speech signals having an echo disturbance.

### A. BACKGROUND OF THE INVENTION

#### 1. Field of the invention

5           The invention lies in the area of measuring the conversational quality of telecommunications systems. More in particular, it concerns measuring the influence of echo disturbances on the subjectively observed conversational quality of a telephone link in a telecommunications system.

#### 2. Prior art

10           Echo is a significant factor in the perception of the conversational quality in an end-to-end telephone link. The influence of said factor on the perception may be quantified by measuring the combination of a pair of parameters known by the indications "talker echo loudness rating" and "round-trip delay", such as defined, e.g., in the Recommendation G.107 of ITU-T (Reference [1]; for more bibliographical details relating to the references, see below under D.). For a non-intrusive measurement for determining such parameters, ITU-T Recommendation P.561 (see Reference [2]) provides recommendations. Said known technique, however, has the following drawbacks. To measure the influence of the echo disturbance, first of all a correlation technique is required to determine the delay of the echo. Furthermore, a specific algorithm is required for estimating the "talker echo loudness rating". Moreover, determining the echo delay is difficult when the echo signals are of a low level, i.e., when they are located near the limit of perceptibility, as often is the case on high-quality speech links. Even if a correct estimate is made of the delay and the loudness, even then the differences in perception of the echo disturbance between different speakers are not capable of being measured, since during the measurement no details of the speech signal can be taken into account. Therefore, the known technique shows up only a moderate correlation between the objective measurement results and the more subjective findings of the speakers, particularly in the event of slight echo disturbances.

35           In fact, in the event of echo on a telephone link between an A subscriber and a B subscriber, a distinction must be made between a so-called talker echo (from A to A) which the talker at the transmission side (A) may experience during conversation, and a listener echo, which the listener at the receiving side (B) may

**THIS PAGE BLANK (USPTO)**

experience while listening. The perception of the influence on the listener of the listener echo, which in fact consists of twice-reflected speech signals, may basically be quantified using a so-called objective measurement technique, such as the Perceptual Speech Quality Measure (PSQM). With said known technique (see, e.g., references [3] and [4]), which models the perceptual properties of human hearing, the quality of a one-way speech link may be predicted. Quantifying the influence of echo on the perception of the talker using said technique, however, is not known and not possible just like that, but it is desirable.

#### B. SUMMARY OF THE INVENTION

The object of the invention is to provide for a method and a device for measuring the influence of echo on the perceptual quality on the part of the talker of a telephone link in a telecommunications network, which both does not possess said drawbacks of the known technique and accommodates said desire.

A further object of the invention is to provide for a telephone-link circuit in which the method and the device are applied.

A method and a device for measuring the conversational quality of a telephone link in a telecommunications network according to the preamble of claim 1 and claim 9, respectively, for whose definition reference [4] was applied, according to the invention are characterised as in claim 1 and claim 9, respectively.

In a preferred embodiment of the invention, the method is characterised as in claim 2, and the device is characterised as in claim 10.

The invention is based on the insight that a talking telephone user simultaneously listens and therefore hears his own speech signal simultaneously with an echo of his speech possibly returning from the headphone of the telephone set. Therefore, the application of an original speech signal and a combined signal, composed of the original speech signal and a corresponding echo signal, as input signals for an objective measurement technique, such as PSQM, may lead to a usable estimate of the perceptual quality, whereas such is not the case if only the original speech signal and the corresponding echo signal are used.

A telephone-link circuit for a telephone link in a telecommunications network, comprising a forward channel and a return channel, and an echo-minimising device included between the forward

**THIS PAGE BLANK (USPTO)**

channel and the return channel, according to the invention is characterised as in claim 14.

Further preferred embodiments of the invention are summarised in further subclaims.

#### C. REFERENCES

- [1] ITU-T Recommendation G.107: The E-model, a computational model for use in transmission planning, December 1998;
- [2] ITU-T Recommendation P.561: In-service, non-intrusive measurement device - voice service measurements, February 1996;
- [3] J.B. Beerends and J.A. Stemerdink, A perceptual speech quality measure based on a psychoacoustic sound representation. J. Audio Eng. Soc. 42:115-123, March 1994;
- [4] ITU-T Recommendation P.861: Objective quality measurement of telephone band (300-3400 Hz) speech codecs, August 1996.

All references are considered to be incorporated into the present application.

#### D. BRIEF DESCRIPTION OF THE DRAWING

The invention will be further explained by means of the description of exemplary embodiments, reference being made to a drawing comprising the following figures:

- FIG. 1 schematically shows a known method for measuring the perceptual quality of a speech signal;
- FIG. 2 schematically shows a telephone link in a telecommunications network;
- FIG. 3 schematically shows a method according to the invention;
- FIG. 4 shows part of a telephone-link circuit in which the invention is applied.

#### E. DESCRIPTION OF EXEMPLARY EMBODIMENTS

FIG. 1 shows a known schematical setup of an application of an objective measurement technique, such as, e.g., the one based on a model of the human hearing and which is usually designated by PSQM, for estimating the perceptual quality of speech links. It comprises a system or telecommunications network under test 10, hereinafter referred to as network 10 for briefness' sake, and a system 12 for the perceptual analysis of speech signals offered, hereinafter designated, for briefness' sake only, by PSQM system 12. A speech signal  $d(t)$  is used, on the one hand, as an input signal of the

**THIS PAGE BLANK (USPTO)**

network 10 and, on the other hand, as a first input signal of the PSQM system 12. An output signal  $d'(t)$  of the network 10, which in fact is the speech signal  $d(t)$  affected by the network 10, is used as a second input signal of the PSQM system 12. An output signal  $p(t)$  of the PSQM system 12 represents an estimate of the perceptual quality of the speech link through the network 10. Since the input end and the output end of a speech link, particularly in the event it runs through a telecommunications network, are remote, for the input signals of the PSQM system use is made in most cases of speech signals stored on data bases. Here, as is customary, speech signal is understood to mean each sound basically perceptible to the human hearing, such as speech and tones. The system or network being tested may of course also be a simulation system, which simulates a telecommunications network. With this known technique, reliable estimates of the perceptual quality are possible.

FIG. 2 schematically shows a telephone link established between an A subscriber and a B subscriber of a telecommunications network 20. Telephone sets 21 and 22 of the A subscriber and the B subscriber, respectively, are connected by way of two-wire connections 23 and 24 and four-wire interfaces, namely, hybrids 25 and 26, to the network 20. Through the network, the established telephone link has a forward channel 27, over which speech signals from the A subscriber are conducted, and a return channel 28, over which speech signals from the B subscriber are conducted. A speech signal  $s(t)$  striking the microphone of the telephone set 21 of the A subscriber, is passed on, by way of the forward channel 27 of the telephone link, to the earphone of telephone set 22, and becomes audible there for the B subscriber as a speech signal  $s'(t)$  affected by the network. Each speech signal on the forward channel 27, particularly due to the presence of said hybrids, generally causes an echo signal  $e(t)$  on the return channel 28 of the telephone link, and this is passed on to the earphone of the telephone set 21, and may therefore disturb the A subscriber there. To restrict such echo effects to a minimum, it is generally customary to include, in telephone-link circuits, echo suppressors or echo cancellors, one at each end of a telephone link. Using the signals present in the forward and return channels, an echo suppressor or cancellor continuously makes an estimate of the echo signal and subtracts it from the signal in the forward channel. Such an estimate, however, cannot always be carried out reliably, which is particularly the case on high-quality speech links, where the echo signals are at the

**THIS PAGE BLANK (USPTO)**



audibility limit. The optimisation routines applied within such echo suppressors or cancellors do not always provide the best result concerning the observation. Should such optimisation routines be based, at least in part, on the method described below with reference to FIG. 3, then an optimum result having a minimally perceptible echo is possible.

FIG. 3 shows a schematic setup according to the invention, for obtaining an estimate of the perceptual quality of a telephone link, for a telephone user when speaking on his own telephone set. In a similar manner as the setup of FIG. 1, the setup of FIG. 3 comprises a system or telecommunications network under test 30, hereinafter for briefness' sake referred to as network 30, and a system 32 for the perceptual analysis of speech signals offered, hereinafter for briefness' sake only designated as PSQM system 32. Any speech signal  $s(t)$  is used, on the one hand, as an input signal of the network 30 and, on the other hand, as first input signal of the PSQM system 32. An echo signal  $e(t)$  obtained from the network 30, which corresponds to the input speech signal  $s(t)$ , is combined, in a combination circuit 34, with the speech signal  $s(t)$  to provide a combined speech signal  $s'(t)$ , which is then used as a second input signal of the PSQM system. An output signal  $q(t)$  of the PSQM system 32 represents an estimate of the perceptual quality of the speech link through the network 30 as it is experienced by the telephone user during speech on his own set. Here, too, use may again be made of signals stored on data bases. These may be obtained, e.g., from the telephone set (by electrical tapping from the microphone circuit and the earphone circuit, respectively) of the A subscriber in the event of an established link during speech silence of the B subscriber. Since, however, the two-wire connection between the telephone subscriber access point and the four-wire interface with the network does not, or hardly, contribute towards the echo signal  $e(t)$  (of course, it does contribute towards the echo signal occurring at the B subscriber), the signals  $s(t)$  and  $e(t)$  may also be tapped off from the forward channel and the return channel near the four-wire interface. This offers the opportunity of a permanent measurement of the perceptual quality in the event of established telephone links. For this purpose, FIG. 4 schematically shows an embodiment.

FIG. 4 shows, in a similar manner as part of FIG. 2, a two-wire connection 41 which, by way of a four-wire interface, in this case hybrid 43, is connected to a telecommunications network 40, and by way of which a established telephone link with a forward channel 44

**THIS PAGE BLANK (USPTO)**

and a return channel 45 through the network may be set up. The line circuit belonging to the telephone link includes an echo cancellor 46. Also included is a PSQM system 42, of which a first input port 47 is coupled to the forward channel 44, and a second input port 48 is coupled to an output port of a signal combiner 49 having two input ports which are coupled to the forward channel 44 and the return channel 45, respectively. An output port 50 of the PSQM system 42 for quality-control purposes may be coupled directly, or by way of a switch 51 (arrows F and G), to a monitoring system (not shown). In addition the output port 50, as shown, may be coupled, by way of the switch 51, to a control input 52 of the echo cancellor 46. The switch 51 is preferably controlled by a control signal given off by a detection circuit 53 (constructed, e.g., as a "double-talk" detection circuit known per se), which is coupled to the return channel for detecting the speech status on the return channel 45, such as, e.g., speech silence on the part of the B subscriber. Thus, the estimated signal becoming available by way of the output port 50 of the PSQM system may be used, on the one hand, for all kinds of quality-control purposes and, on the other hand, may be used directly in echo-minimising equipment.

In the most simple embodiments, the combination circuit 34 and the signal combiner 49 are signal adders. When applying the method and the device in practice, in the signal combiner carrying out the adding function (addition) is preferably preceded by the so-called "inversely filtering" of one of the signal components. The inverse filter applied there generates a linear estimate of the echo path, and to a major degree contributes towards achieving a high correlation between an objective measurement and a subjective observation.

**(THIS PAGE BLANK (USPTO))**

F. CLAIMS

1. Method for measuring the conversational quality of a telephone link in a telecommunications network, comprising the following steps:

- 5       - receiving a first speech signal transmitted or to be transmitted at the transmission side of the telephone link,
- receiving a second speech signal, which is a function of the first speech signal affected in the telecommunications network,
- 10       - using an objective measurement technique, processing the first and the second speech signal, an output signal being produced which represents an estimated value concerning the conversational quality,

characterised in that

- the second speech signal is generated by a signal combination of  
15       an echo signal taken off at the transmission side of the telephone link during the transmission of the first speech signal, and the first speech signal.

2. Method according to claim 1, characterised in that the signal  
20 combination comprises a signal addition of the echo signal and the first speech signal.

3. Method according to claim 1, characterised in that the signal  
25 combination comprises a signal addition preceded by an inverse filtering of either the echo signal, or the first speech signal.

4. Method according to claim 1, 2 or 3, characterised in that the  
echo signal originates from a return channel of the telephone link.

30   5. Method according to claim 1, 2, 3 or 4, characterised in that the first speech signal, and either the second speech signal or the echo signal are signals laid down in a data base.

35   6. Method according to claim 1, 2, 3 or 4, characterised in that the first speech signal and the echo signal are taken off from a transmitting side, or from a forward channel and a return channel of an established telephone link, respectively.

40   7. Method according to claim 6, characterised in that the output signal of the objective measurement is fed to a control input of an echo-minimising device included in the established telephone link.

**THIS PAGE BLANK (USPTO)**

8. Method according to claim 6 or 7, characterised in that the output signal of the objective measurement is fed to a monitoring system.

9. Device for measuring the conversational quality of a telephone link in a telecommunications network, comprising:

- a first input port for receiving a first speech signal transmitted or to be transmitted at the transmission side of the telephone link,
- a second input port for receiving a second speech signal, which is a function of the first speech signal affected in the telecommunications network,
- an output port for an output signal, and
- processing means for, using an objective measurement technique, processing the first and second speech signals received by way of the input ports, the output signal being produced which represents an estimated value concerning the conversational quality,

characterised in that

- the device additionally comprises signal-combination means for combining an echo signal taken off at the transmission side of the telephone link during the transmission of the first speech signal, and the first speech signal, and generating the second speech signal.

10. Device according to claim 9, characterised in that the signal-combination means comprise a signal adder.

11. Device according to claim 9 or 10, characterised in that the first and the second input ports are coupled to a data base of speech signals, on which the first speech signal, and either the second speech signal or the echo signal, are laid down.

12. Device according to claim 9, 10 or 11, characterised in that the signal-combination means are provided with a first and a second signal input, which are coupled to the forward channel and the return channel of an established telephone link, respectively, and that the first input port is coupled to the forward channel, and the second input port is coupled to the signal output of the signal-combination means.

**THIS PAGE BLANK (USPTO)**



13. Device according to claim 12, characterised in that the output port is coupled to a control input of an echo-minimising device included in the established telephone link.

5

14. Telephone-link circuit for a telephone link in a telecommunications network, comprising a forward channel and a return channel, and an echo-minimising device included between the forward channel and the return channel,

10

characterised in that the telephone-link circuit further comprises

- an objective measurement device provided with a first and a second input port, and an output port, for processing a first speech signal received on the first input port, and with a second speech signal received on the second input port, and for producing an output signal on the output port, which represents an estimated value concerning the conversational quality of the telephone link, the first input port of the measurement device being coupled to the forward channel of the telephone link, and
- a signal combiner provided with a first and a second signal input, which are coupled to the forward channel and the return channel of a telephone link, respectively, and of a signal output which is coupled to the second input port.

15

20

25

15. Telephone-link circuit according to claim 14, characterised in that the output port has a signal coupling with a monitoring system.

16. Telephone-link circuit according to claim 14 or 15, characterised in that the output port of the measurement device has a signal coupling with a control input of the echo-minimising device.

30

17. Telephone-link circuit according to claim 14 or 15, characterised in that there is further provided for a detection device for detecting the speech status over the established telephone link, and for a switch included in the signal coupling, which is driven by the detection device.

35

**THIS PAGE BLANK (USPTO)**

1/2

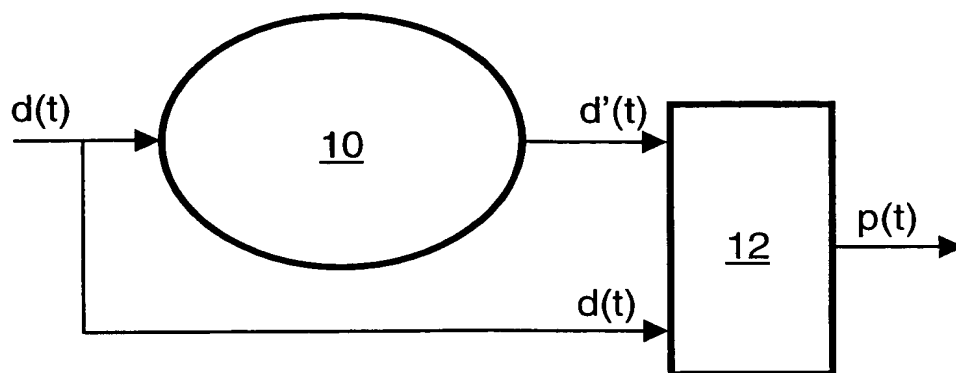


FIG. 1

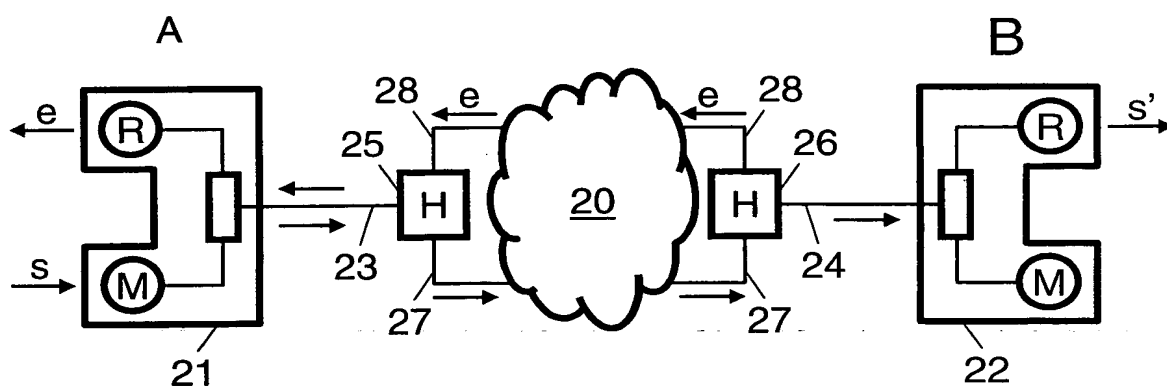


FIG. 2

**THIS PAGE BLANK (USPTO)**

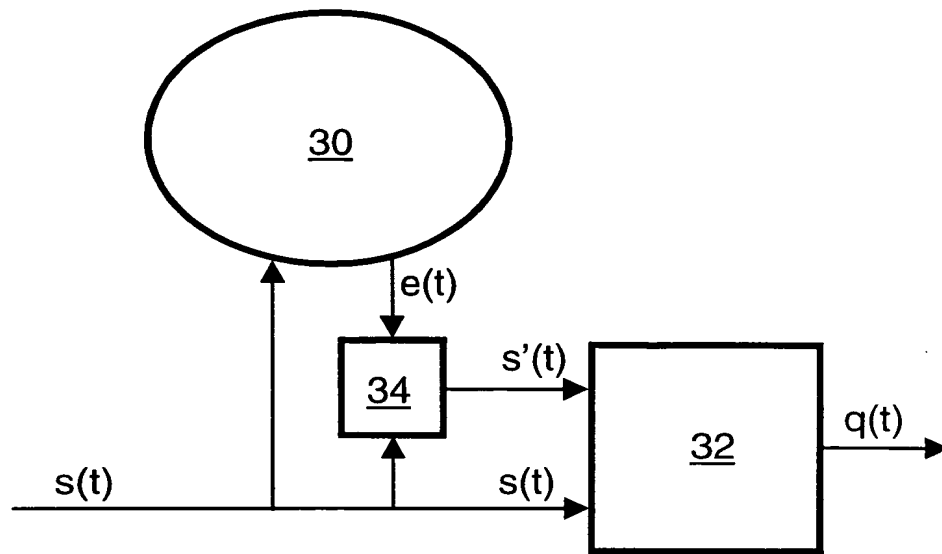


FIG. 3

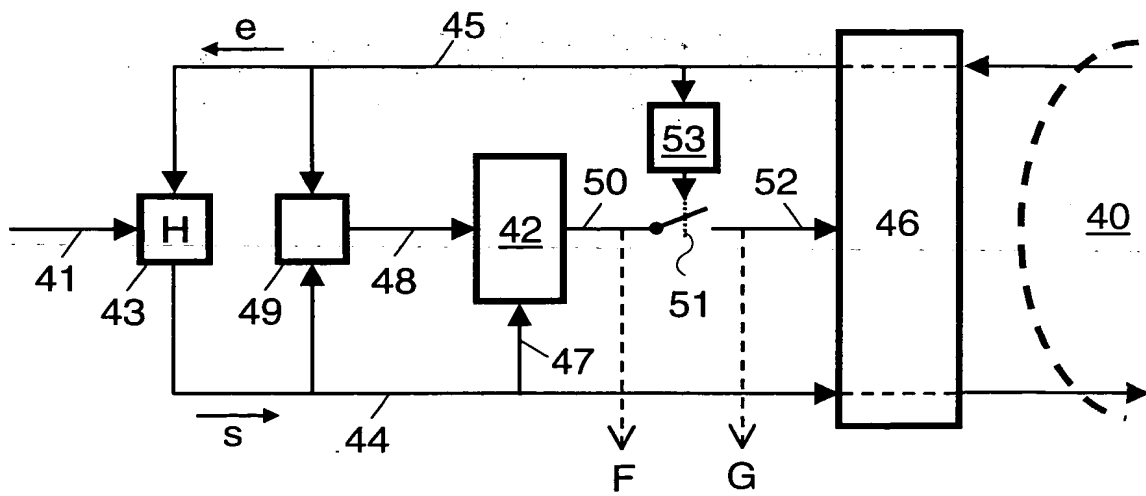


FIG. 4

**THIS PAGE BLANK (USPTO)**